Role of Data Mining on Educational Data Base (U-DISE)

Sheeraz Ahmad Peerzada[1] , Zubair Ahmad[2] , Tahir Ahmad[3] , Dr.Jitendra Seethalani[4]

[1,2,3,4]Department of Computer Science,

Sri Satya Sai University of Technology and Medical Sciences,

Sehore, MP, (India)

ABSTRACT

Schools either Government, private or public enroll lakhs of students every year. The Department of School Education in India collect the information from students , as well as from institutions and store it in a universal data base called unified district information system (U-DISE).The existing system maintains the students ,as well as institution information in the form of numerical values and it just stores and retrieves information that it contains. The system has no intelligence of its own to analyze the data and make decisions. Mining of such data yields stimulating information that serves its handlers well. The main focus of this paper is to apply Educational Data Mining on Unified district information system (U-DISE) in order to derive new patterns that will help in decision making.

Keywords-Data Mining, Educational Data Base (U-DISE), Educational Data Mining,U-DISE Indicators.

I.INTRODUCTION

Data mining refers to extracting or "mining" knowledge from large amount of data .Data mining is a process of discovering interesting knowledge from large amount of data from stored either in data bases, data warehouse or other information repositories.

In recent years, there has been enormous growth of data related to student as well as Educational Institutions. The Department of Education is collecting information on yearly basis in India, and last reference date is 30th of September every year. The said data is stored in a nationalized Data Base called Unified District Information System for Education (U-DISE).

Data Mining can be used in educational field to extend learning process and overall performance and scenario of education system.

Unified District information system for Education (U-DISE) it is developed by National University of Education Planning and administration (NUPA) for successful implementation of Sarva Shiksha Abhiyan (SSA).

U-DISE software is operational all over India. The Central Government has launched U-DISE for collection statistical data, every school must fill information on U-DISE data capture format yearly the information like

students, teachers, and class rooms, facilities, working hours, student hours and days of working . This information is valid for schemes like MDM, SSA, RMSA, ICT.

II. RELATED WORK

J K Jothi and K Venkatalakshmi conducted the student's performance analysis on the graduate students 'data collected from the Villupuram Colleg of Engineering and Technology [1].

Rashi Bansal,Akansha Mishra and Dr.Shailendra Narayan Singh conducted research on Mining of Educational Data for analyzing students'overall performance the main focus of their research was application areas of EDM,where we can apply e-learning[2].

Yadav, Pal have shown the data mining techniques for predicting the results of students. The current results have been analyzed and then comparative analysis of the results has been done. This shows that the prediction has helped the average students to perform better in the final examinations [3].

Rashi Bansal,Akansha Mishra and Dr.Shailendra Narayan Singh conducted research on Mining of Educational Data Mining and learning Analysis the main focus of their research was application areas of like clustering[4].

III. EDUCATIONAL DATA MINING

Educational Data Mining (EDM) is an emerging discipline, concerned with developing methods for exploring the unique types of data that come from educational settings, and using those methods to better understand students, and the settings which they learn in.[6].

Educational Data Mining (EDM) describes a research field concerned with the application of data mining, machine learning and statistics to information generated from educational settings (e.g., universities and intelligent tutoring systems)[7].

Educational Data Mining focuses on developing new tools and customizes the traditional data mining algorithms for analysis and discovering data patterns. EDM develops methods and applies techniques from statistics, machine learning, and data mining to analyze the data stored in Educational Data Base.

 Educational data mining researchers [8] [9] took the following points as their research goal-

1. To predict behavior of future learning for students.

2. To improve or discover models of domain that contains the learning content with other sequences for optimization.

3. To study impact of various types of facilities and their impact on learning process.

4. To advancing Knowledge about booth learning and the learning by building computational models.

Goals of Educational Data Mining:

1. Predicting student's future learning behavior by creating student models on the different U-DISE data indicators.

2. Discovering the models that characterize the content and curriculum to be taught in schools.

3. Studying the comparison of different Data mining algorithms and select the one optimal for particular indicator of U-DISE.

4. Advancing scientific knowledge about learning and learners through building computational models that incorporate models of the student, the domain, and the software's pedagogy.

IV.U-DISE, INDICATORS AND APPLICATIONS OF EDUCATIONAL DATA MINING

For each indicator different Data Mining algorithms may be applied and the algorithm with optimal results may be used for improving overall performance of students and look in to week areas of students as well as institutions.

4.1. School Based Indicators.

4.1.1The information related to School such as number of Schools, PS, MS,UPS,HS,HSS and Army Schools, JNV's and Maktabs.

4.1.2 The type of school such as Boys only, Girls only or coeducational.

4.1.3 Wither the school is Government, private, public, aided or un aided.

4.1.4 Year of establishment of school.

4.1.5 School Building type such as Kacha, packa etc.

4.1.6 Infrastructure, number of class rooms.

4.1.7 The percentage of single class room schools

4.1.8 The student class room ratio.

4.2. Enrolment Related Indicators

4.2.1 With the increased coverage of schools under U-DISE this indicator plays important role related to enrolment in these schools.

4.2.2 Enrolment booth at Primary and Secondary level.

4.2.3 The general enrolment ratio (GER) and net enrolment ratio (NER) .

4.2.4 The girl enrolment.

4.2.5 Enrolment SC, ST and OBC category wise.

4.3. Teacher Related Indicators

4.3.1 The information related to teachers like availability of teachers in schools.

4.3.2 Total number of teachers in schools.

4.3.3 Total number of teachers for primary, Upper primary and secondary.

4.3.4 Percentage of male, female teachers.

4.3.5 Pupil teacher ratio (PTR).

4.4. Facility Related Indicators

4.4.1 All facilities provided by school like number of schools having electricity.

4.4.2 Number of schools having primary sections attached.

4.4.3 Number of schools having water facility.

4.4.4 Number of schools having CAL and ICT facility.

4.4.5 Number of Schools having RAMPS, Kitchen shed, Boundary walls.

4.4.6 Whether MDM is served or not by School.

The Data to be stored in Universal data base (U-DISE) has to be checked at different levels manually such as by school Teacher, Head Teacher, Zonal/Block Officer then the operator makes data entry after that the data is stored in a data base, different consistency checks present in the software are also implemented at Cluster/Zone/District/State level. The U-DISE store all information in the form of numerical values these values had unique meaning such as category of teacher, facility present in school or not in numerical form such as 1 for yes, and 2 for no etc.

Statistical figures are always available in the form of reports. U-DISE data is consistent.

V.CONCLUSION

The increased use of Technology in education is generating large amount of data every day from students as well as from educational institutions which has become a target for many researchers around the world. This paper discusses many Educational Data Mining applications on Unified District Information System (U-DISE).By knowing all these applications it is easier to implement them. To know all these indicators it became easier to make policies and focus on week areas. I am going to implement these applications in my research so that platform made for its real implementation.

VI.ACKNOWLEDGEMENTS

REFERENCES

[1] J.K Jothi and K.Venkatalakshmi, "Intellectualt performance analysis of students by using data mining techniques.", International Journal of Innovative Research in Science,Engineering and Technology,vol 3,special iss 3, March 2014.

[2] Rashi Bansal,Akasha Mishra and Dr.Shilendra Narayan , "Mining of educational data for analysisng student's over all performance ", IEEE Journal.

[3] Yadav, Pal, " Data Mining: A Prediction for Performance Improvement of engineering Students using Classification", *World of Computer Science and Information Technology Journal (WCSIT) , 2012* [5] Edin Osmanbegović , Mirza Suljić "Data Mining Approach For Predicting Student Performance**,** *Economic Review – Journal of Economics and Business"*, May 2012.

[4] Rashi Bansal,Akasha Mishra and Dr.Shilendra Narayan , "Educational Data Mining and Learning Analysis ", IEEE Journal 2017.

[6] Wikipidia.

[7] Wikipidia.

[8] Baker ,r,s,j,d Data Mining for Education. In :McGaw,Peterson,Baker,E.(eds),Oxford (2010).

[9] Witten, Frank and I.H(1999). Morgan Kauffman,San Fracncisco,CA.